

ABSTRACT

A perceptual audio coder is disclosed for encoding audio signals, such as speech or music, with different spectral and temporal resolutions for redundancy reduction and irrelevancy reduction. The disclosed perceptual audio coder separates the psychoacoustic model (irrelevancy reduction) from the redundancy reduction, to the extent possible. The audio signal is initially spectrally shaped using a prefilter controlled by a psychoacoustic model. The prefilter output samples are thereafter quantized and coded to minimize the mean square error (MSE) across the spectrum. The disclosed perceptual audio coder can use fixed quantizer step-sizes, since spectral shaping is performed by the pre-filter prior to quantization and coding. The disclosed pre-filter and post-filter support the appropriate frequency dependent temporal and spectral resolution for irrelevancy reduction. A filter structure based on a frequency-warping technique is used that allows filter design based on a non-linear frequency scale. The characteristics of the pre-filter may be adapted to the masked thresholds (as generated by the psychoacoustic model), using techniques known from speech coding, where linear-predictive coefficient (LPC) filter parameters are used to model the spectral envelope of the speech signal. Likewise, the filter coefficients may be efficiently transmitted to the decoder for use by the post-filter using well-established techniques from speech coding, such as an LSP (line spectral pairs) representation, temporal interpolation, or vector quantization.

1200-410.app